



Access Control for NCAR Data Portals

A report on work in progress
about the future
of the NCAR Community Data Portal

Luca Cinquini

GO-ESSP Workshop, 6-8 June 2006

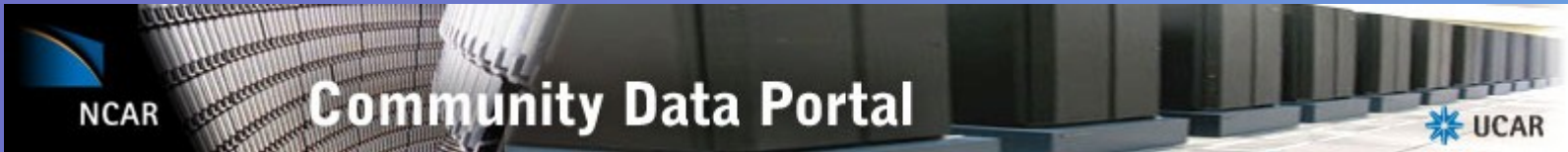
Acknowledgments: Don Middleton, Rob Markel, Mike Burek



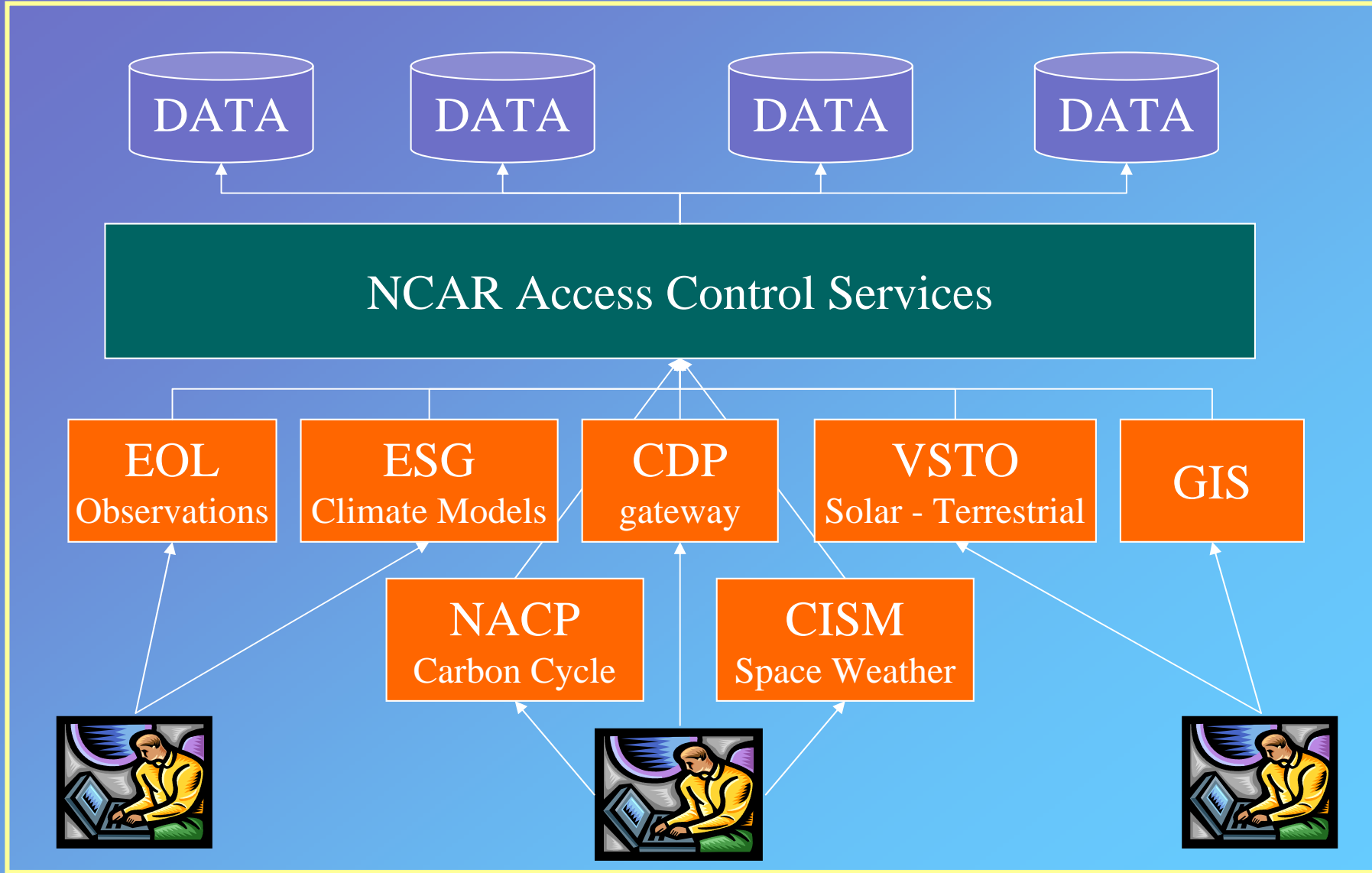
NCAR Scientific Computing Division
Supercomputing • Communications • Data

Introduction

- Web-based access to data and services is extremely popular and effective – one of the preferred ways through which scientists and other users find and download their data
- Several NCAR data portals already exist, are under development or planned: CDP, ESG, GIS, GridBGC, VSTO, EOL, DSS, JOSS, ...
- Necessity to coordinate a unified NCAR strategy for web-based data access
 - Avoid multiple logins, registration
 - Present consistent interface and suite of services to users
 - Avoid duplication of effort



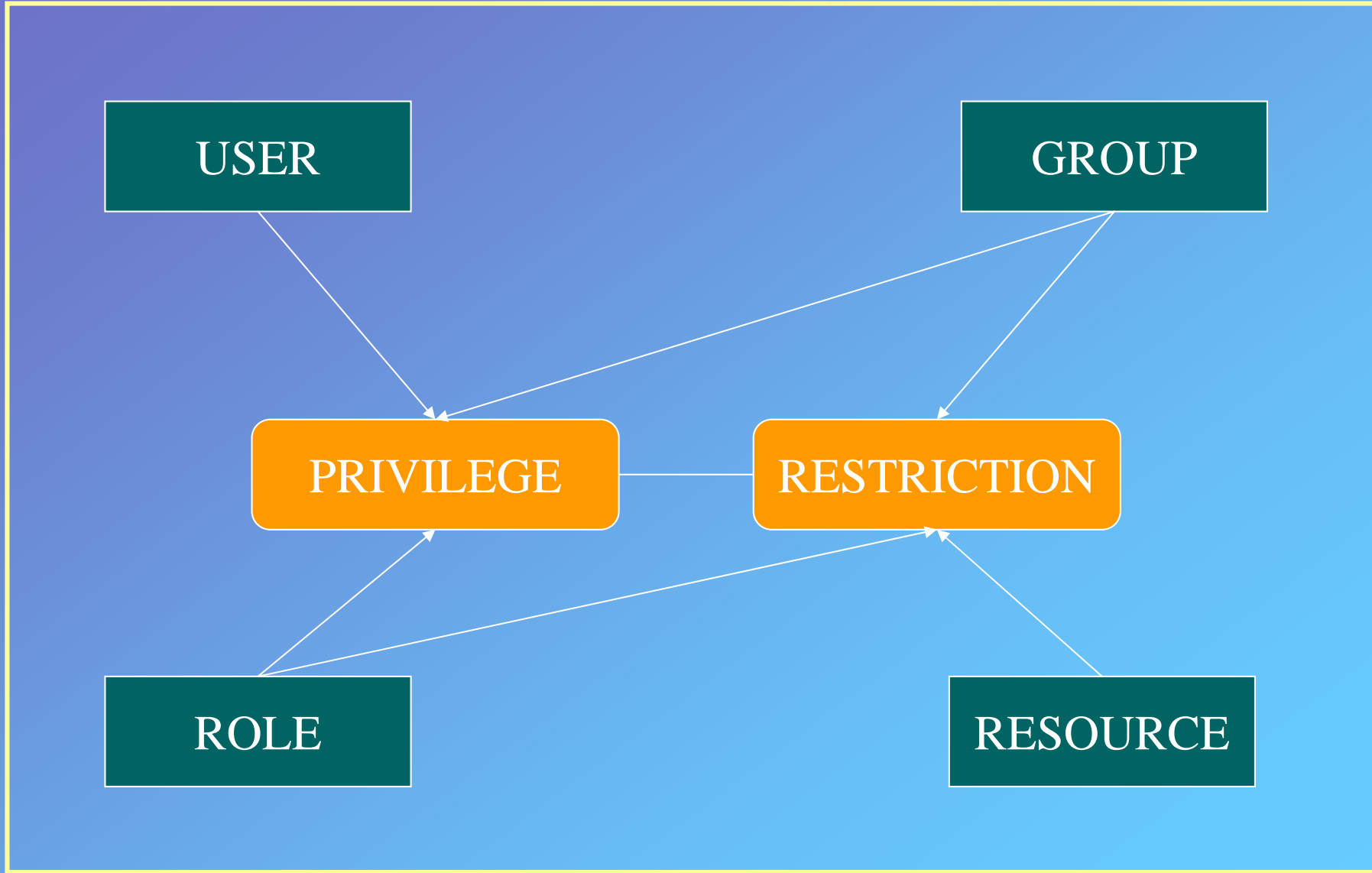
- Proposal: system of federated NCAR data portals that share technology, data and metadata:
 - CDP: central gateway to all NCAR data holdings, offers basic services:
 - Generic search & discovery
 - Catalogs browsing
 - Metadata exchange with partner institutions
 - Multiple discipline-specific portals
 - Increased discipline-specific functionality
 - Decentralized management of data holdings and services
 - Allow “branding”
- Need common set of services used by all portals to establish access control to shared, distributed resources

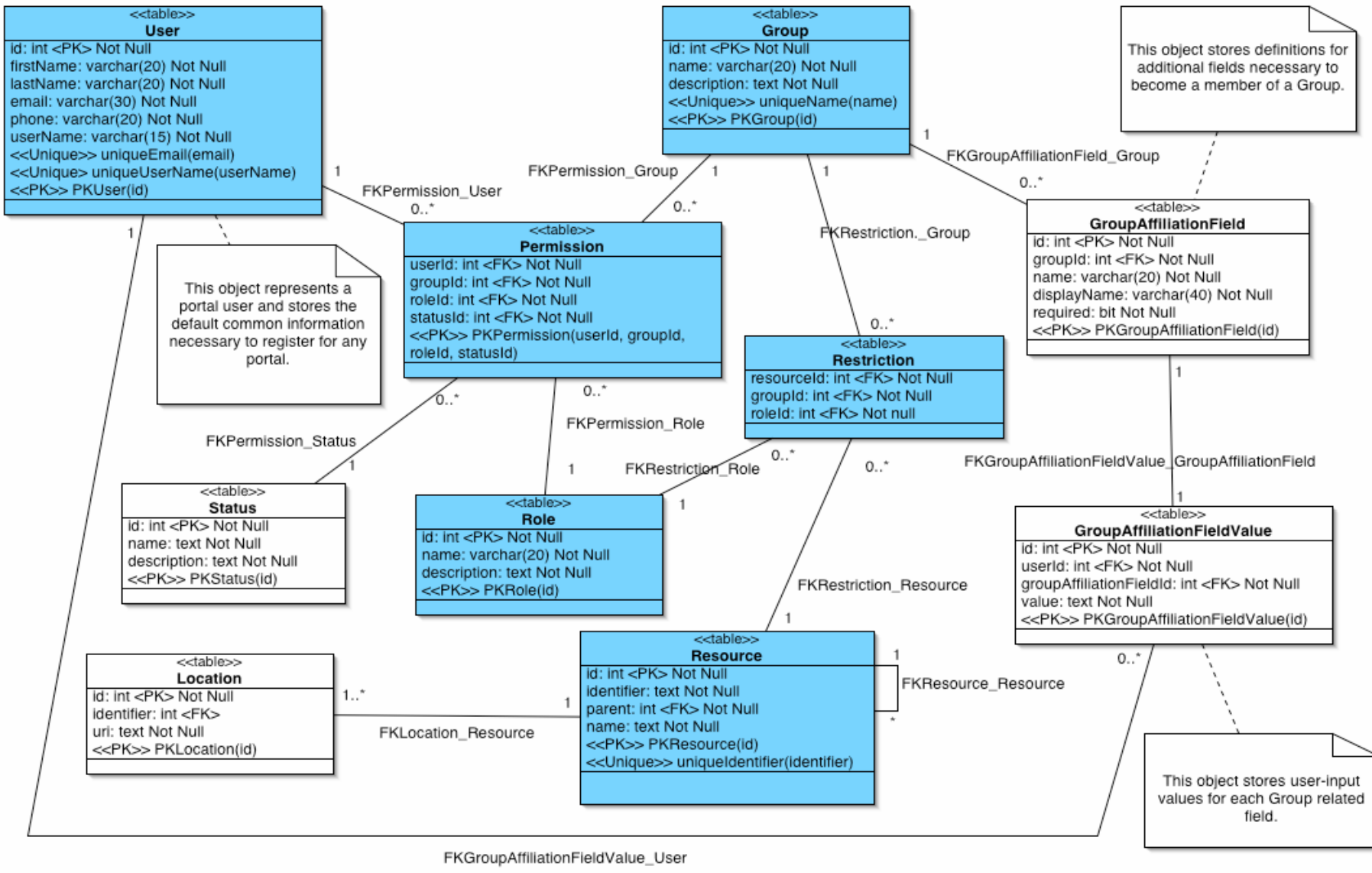


Access Control Model

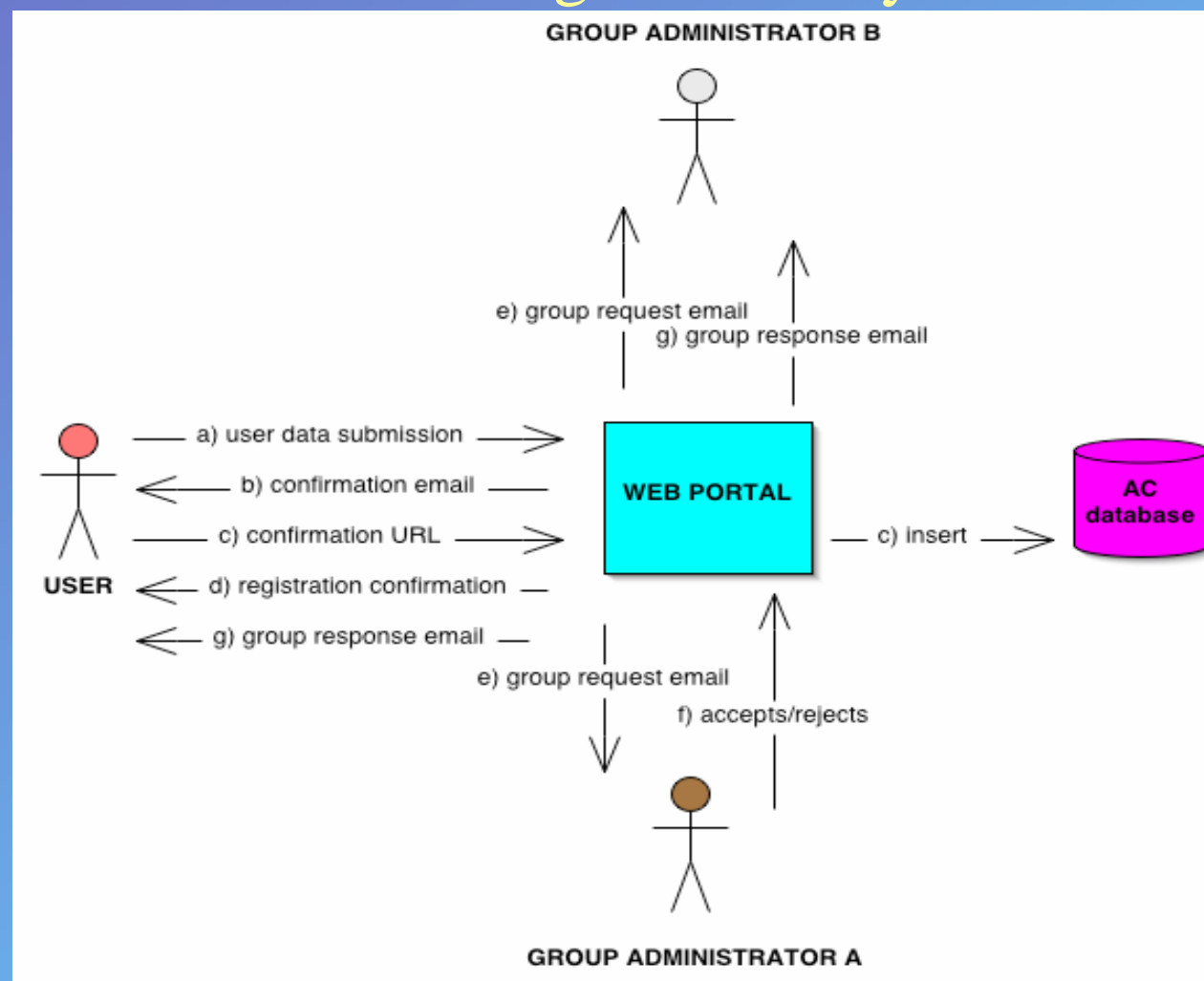
Objects: users, groups, roles, resources (= data or services)

- Group: affiliation of a user to a VO (“ccsm”, “vsto”, etc.)
- Role: level of capability within group (“admin”, “read”, etc)
- User is assigned one or more privileges: (user, group, role)
 - (“Johnny”, “ccsm”, “read”) or (“jimmy”, “gis”, “admin”)
- Resource may be subject to one or more restrictions: (resource, group, role)
 - (“/disk/data/ccsm”, “ccsm”, “write”)
 - (“http://host:port/data/ccsm/”, “user”, “read”)
 - Note: if null, restrictions are inherited from parent resource
- Authorization: match user privileges to resource restrictions

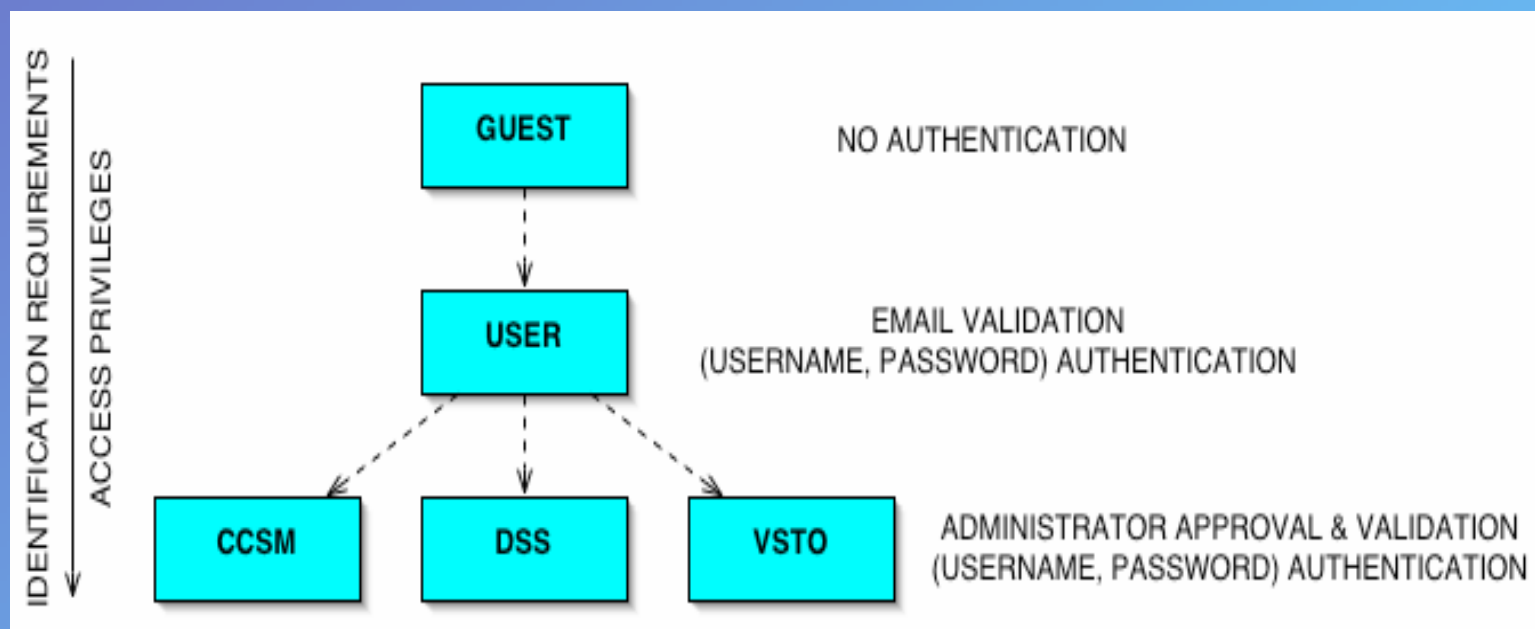




Web-based Registration System



Groups Hierarchy



Identification of Data Resources

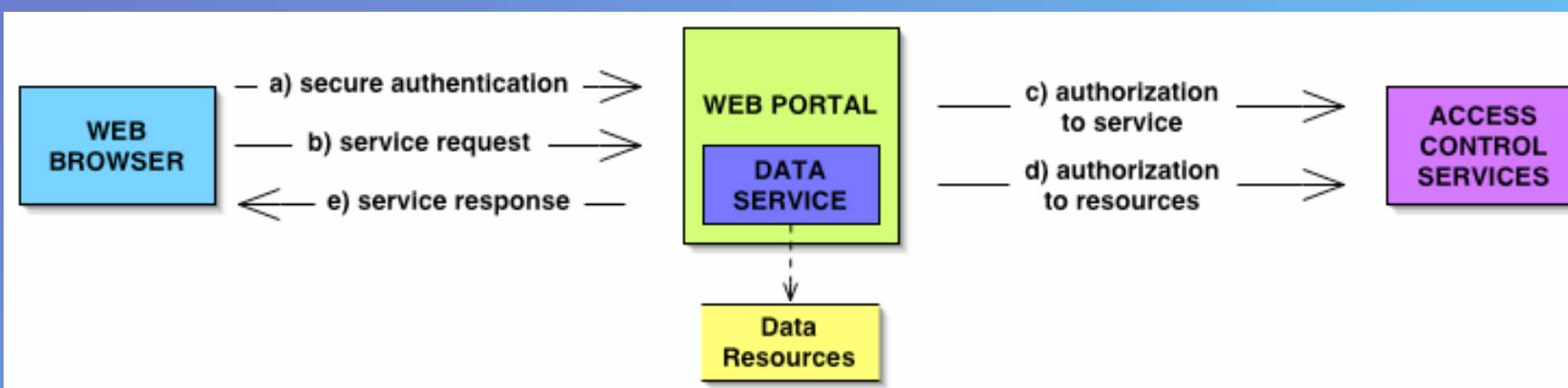
Resources must be uniquely identifiable so that restrictions may be set up and looked up

- By logical id (ex: `ucar.cgd.ccsmb30.004.atm.file1.nc`)
 - + Clear distinction between logical resource and physical location
 - + Automatic sharing of restrictions among replicas of a logical resource
 - Must store logical id - physical location mapping for each replica
 - Must store restriction on each resource OR restriction on some ancestor and full resources hierarchy
- By URI (ex: `http://host:/data/ccsm.b30.004.atm.file1.nc`)
 - + Resources hierarchy is implicit (contained in directory structure)
 - + No separate storage of logical id
 - + Restrictions may be reduced to a minimum (i.e. imposed on directories)
 - Restrictions must be imposed separately for replicas (and multiple access protocols)

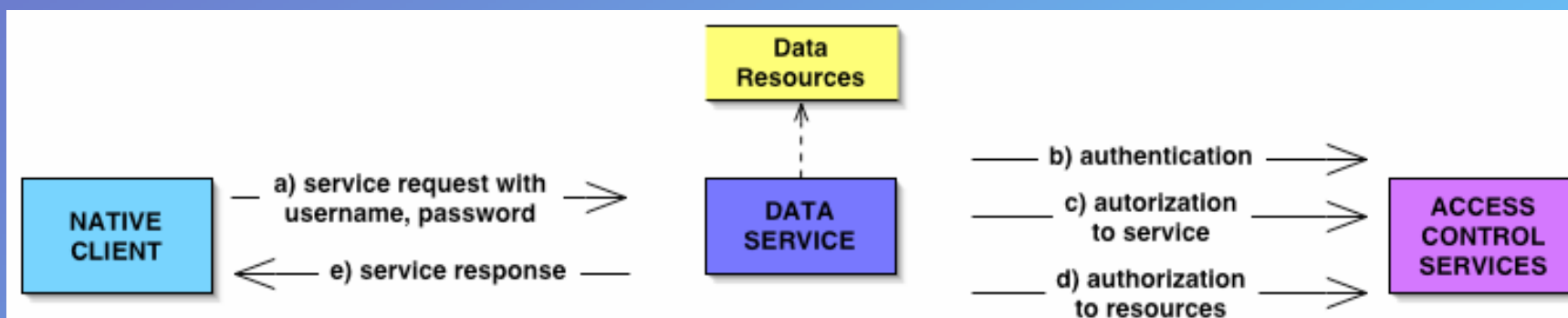
Access Control to Data Use Cases

1. Portal-executed authorization
 - Service is embedded within portal
 - Portal contacts Access Control Services to establish authorization before invoking service
 - Service is agnostic to authorization process
2. Service-executed authorization
 - Native client makes request to standalone service via secure protocol (including authentication information)
 - Data service contacts Access Control Services to establish authorization
 - Data service needs to be instrumented as Access Control Services client
3. Portal-delegated authorization
 - Client-service communication is insecure (example: http download)
 - Request is sent first to portal which executes authorization
 - Request is then redirected to service
 - Service validates request

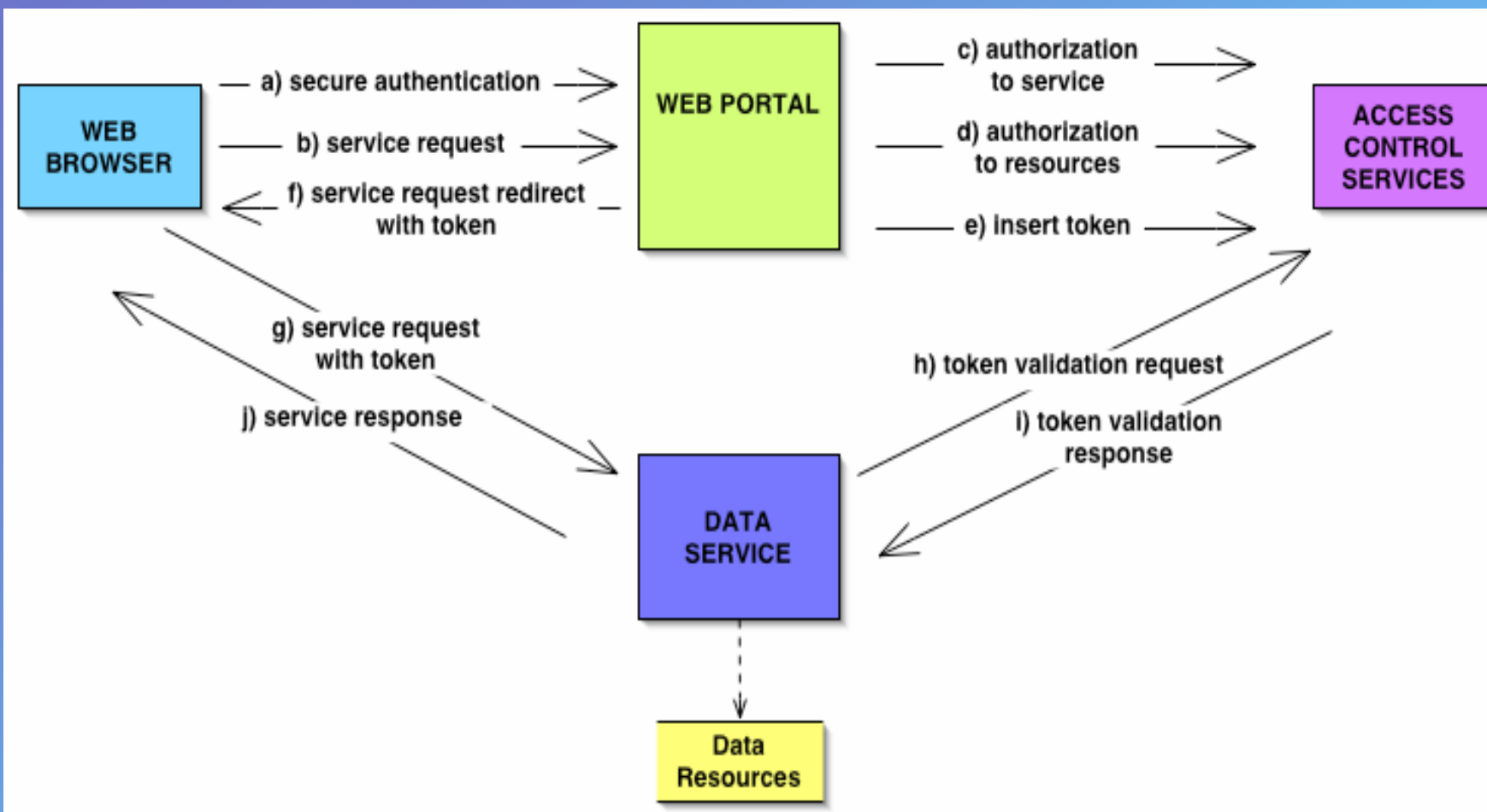
Use Case 1: Portal-Executed Authorization

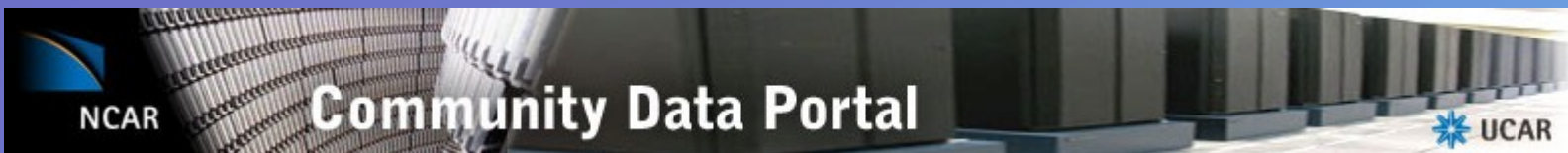


Use Case 2: Service-Executed Authorization



Use Case 3: Portal-Delegated Authorization





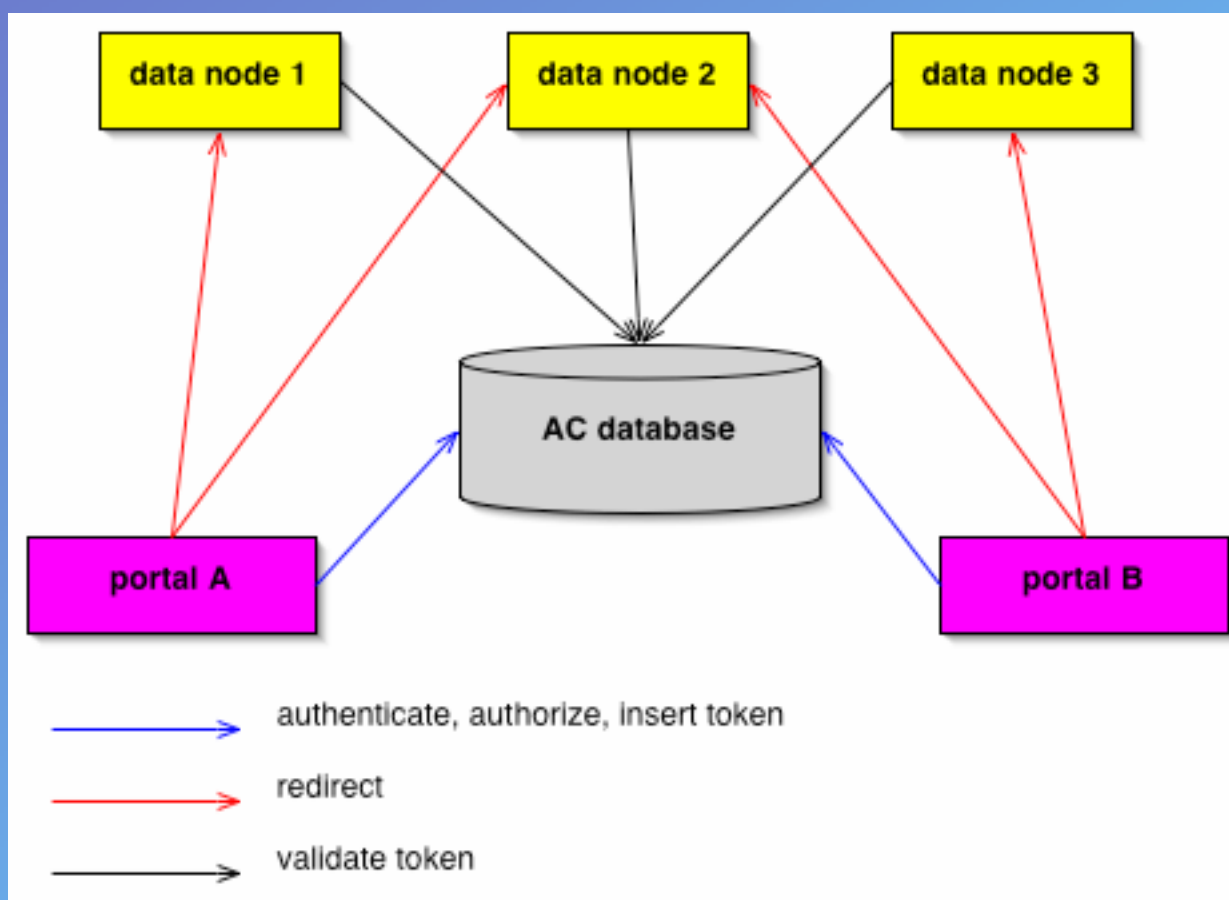
Demo

restricted http download



NCAR Scientific Computing Division
Supercomputing • Communications • Data

LAHFS: Lightweight Authorized Http File Server





Summary

- NCAR Community Data Portal is undergoing evolutionary phase: from single portal to gateway of federated, discipline-specific portals
- Access Control mechanism is a cornerstone of federation model
 - Some system components already developed and deployed
 - Software needs to be upgraded and integrated into deployable, production-level package
- Feedback and suggestions are welcome
- Progress report at next year GO-ESSP meeting